Claims:

1.   A method for monitoring information content carried in a medium, the method comprising:

5   monitoring said medium for said information;

seeking elementary information units within objects of said information being monitored in said medium;

identifying said elementary information units; and

deducing information about the content of said information objects from

10   identification of said elementary information units found within said objects.

2.   A method according to claim 1, wherein said medium comprises at least one of the following:

a distribution channel; and

15   a storage medium.

3.   A method according to claim 1, wherein said information objects comprise at least one simple information object, said simple information object comprising one of the following:

20   an elementary information unit;

a set of elementary information units; and

an ordered set of elementary information units.

4.   A method according to claim 1, wherein said elementary

25   information units comprise at least one of the following:

a sentence; a sequences of words; a word; a sequence of characters; a character; a sequence of numbers; a number; a sequence of digits; a digit; a vector; a curve; a pixel; a block of pixels; an audio frame; a musical note; a musical bar; a visual object; a sequence of video frames; a sequence of musical

30   notes; a sequence of musical bars; and a video frame.

5.   A method according to claim 1, further comprising assigning elementary information units identifiers to elementary information units after identification.

51

6. A method according to claim 5, wherein said elementary information unit identifiers are utilized in said deducing.

7. A method according to claim 1, wherein said information object identification is carried out on an instance of said information object, said information object instance being said information object in a specific format.

8. A method according to claim 7, wherein said format comprise at least one of the following:

jpeg image; gif image; Word document format; Lotus notes format; mpeg format; text format; rich text format; Unicode text format; multi byte text encoding format; formatted text format; ASCII text format; HTML; XML; PDF; postscript; MS-Excel spreadsheet; MS-Excel drawing; MS-Visio drawing; Photoshop drawing; AutoCAD drawing format; and CAD drawing format.

9. A method according to claim 5, wherein said elementary information unit identifiers are determined by the content of said elementary information units which they are assigned to.

10. A method according to claim 9, wherein said elementary information unit identifiers are solely determined by said content.

11. A method according to claim 5, wherein said elementary information units identifiers are at least partly determined by locations within an information object of respective elementary information units to which they are assigned.

12. A method according to claim 5, wherein said elementary information units identifiers are at least partly determined by the content of an elementary information unit in proximity to said elementary information units to which they are assigned.

13.   A method according to claim 5, comprising storing said elementary information units identifiers in a database.

14.   A method according to claim 13, further comprising using said elementary information units identifiers stored in said database for identifying at least one further, unidentified, information object.

15.   A method according to claim 13, further comprising using said elementary information units identifiers stored in said database for comparing information objects.

16.   A method according to claim 5, comprising storing only some of said elementary information units identifiers in a database.

17.   A method according to claim 16, wherein said storing of only some of said elementary information units identifiers in a database is to achieve at least one of the following:

reduce storage cost;

increase efficiency of assigning of said elementary information units identifiers to said elementary information units by only performing said assignment for elementary information units identifiers that are stored in said database; and

increase the efficiency of searching for said elementary information units identifiers in said database.

18.   A method according to claim 16, wherein said storage of only some of said elementary information units identifiers in a database is done in a manner that ensures that any area of a given size in said information object contains a predetermined minimum number of said stored elementary information units.

19.   A method according to claim 18, wherein said given size is dependent on properties of a respective information object.

20. A method according to claim 19, wherein said properties of said information object comprise at least one of the following:

importance; size; confidentiality level; and format.

21. A method according to claim 18, wherein said minimum number is dependent on properties of said information object.

22. A method according to claim 21, wherein said properties of said information object comprise at least one of the following:

importance; size; confidentiality level; and format.

23. A method according to claim 3, wherein said information objects comprise at least one compound information object, said compound information object comprising at least one of the following:

a simple information object; a compound information object; an ordered set of compound information objects; an ordered set of simple information objects; and an ordered set of compound and simple information objects.

24. A method according to claim 1, wherein said information comprises at least one of the following:

numeric data; spreadsheet data; numeric spreadsheet data; textual spreadsheet data; word processor data; textual data; hyper text data; audio data; visual data; multimedia data; binary data; raw data; database data; video data; drawing data; chart data; picture data; and image data.

25. A method according to claim 1, wherein monitoring is done in at least one of the following:

Firewall; Web server; Web proxy; HTTP proxy; HTTP server; SMTP gateway; SMTP server;

Fax server; SOCKS proxy; Sniffer; Server; WAN gateway; proxy; Router; Mail server; file server; client; file system; gateway; router; application; operating system; database; database accessing utility; database accessing server; Internal mail server; External mail server; Message board; NNTP server; and an IRC server.

26.   A method according to claim 1, wherein monitoring is carried out on at least one of the following traffic types:

Instant messaging; IP; HTTP; Mail; TCP; UDP; Web; Streaming; Chat; IRC; computer network; LAN; WAN; VPN; POP3; MAPI; FTP; NNTP; File transfer; IMAP; SMTP; and Fax.

27.   A method according to claim 1, wherein monitoring is done by at least one of the following:

buffering; caching; forwarding; sniffing; and relaying.

28.   A method according to claim 1, wherein monitoring comprises at least one of the following:

blocking traffic; altering traffic; and altering traffic such as to invalidate said traffic.

29.   A method according to claim 1, comprising carrying out said monitoring at a proxy.

30.   A method according to claim 29, comprising routing traffic to be monitored to said proxy.

31.   A method according to claim 29, comprising blocking any traffic requiring monitoring which manages to bypass said proxy.

32.   A method according to claim 31, comprising using a firewall to carry out said blocking.

33.   A method according to claim 29, wherein said proxy is a SOCKS proxy.

34.   A method according to claim 29, wherein said proxy is an HTTP proxy.

35.   A method according to claim 1, comprising monitoring instant messaging traffic.

36.   A method according to claim 35, comprising monitoring file distribution controlled by said instant messaging traffic.

37.   A method according to claim 36, comprising altering said instant messaging traffic controlling said file distribution, thereby to facilitate capturing said file distribution.

38.   A method according to claim 1, comprising using said deducing to attach to said information object an information object policy, said policy comprising at least one of the following:
        an allowed distribution of said information object;
        a restriction on distribution of said information object;
        an allowed storage of said information object;
        a restriction on storage of said information object;
        an action to be taken as a reaction to an event;
        an allowed usage of said information object; and
        a restriction on usage of said information object.

39.   A method according to claim 38, wherein said information object policy comprises at least one action to be taken as a reaction to an event, and wherein said action comprises at least one of the following:
        preventing distribution of said information object; preventing storage of said information object; preventing usage of said information object; reporting distribution of said information object; reporting storage of said information object; reporting usage of said information object; reporting; alerting about distribution of said information object; alerting storage of said information object; alerting usage of said information object; alerting; logging distribution of said information object; logging storage of said information object; logging usage of said information object; logging; notifying about distribution of said information object; notifying about storage of said information object; notifying about usage of said information object; notifying; notifying to an administrator;

notifying to a manager; notifying to a recipient; notifying to a sender; notifying to an owner of said information object; quarantine; alerting an administrator; alerting a manager; alerting a recipient; alerting a sender; alerting an owner of said information object; reporting to an administrator; reporting to a manager; reporting to a recipient; reporting to a sender; reporting to an owner of said information object; encrypting said information object; changing said information object; replacing said information object; and utilizing digital rights management technology on said information object.

40.   A method according to claim 38, wherein said information object policy comprises at least one action to be taken as a reaction to an event, and wherein said event comprises at least one of the following:

attempted distribution of said information object; attempted storage of said information object;

attempted usage of said information object; distribution of said information object; storage of said information object; and usage of said information object.

41.   A method according to claim 38, wherein said information object usage comprises at least one of the following:

copying an excerpt; editing; copying to clipboard; copying an excerpt to clipboard; changing format; changing encoding; encryption; decryption; changing digital management; opening by an application; and printing.

42.   A method according to claim 38, wherein said information object policy comprises placing a substantially imperceptible marking in said information object, said marking comprising information content, and said method comprising placing said marking, when indicated by said policy, before allowing at least one of the following:

storage of said information object; usage of said information object; and distribution of said information object.

43.   A method according to claim 42, wherein said information content for storage in said marking comprises at least one of the following:

the identity of said information object;

the identity of a user performing the action in respect to said information object;

the identity of a user authorizing the action in respect to said information object;

the identity of a user overriding policy and approving the action in respect to said information object; and

the identity of a user requesting the action in respect to said information object.

44.   A method according to claim 38, wherein said information object policy further comprises changing said information object by at least one of the following:

deleting part of said information object; replacing part of said information object; and inserting an additional part to said information object before allowing at least one of the following actions:

storage of said information object; usage of said information object; and distribution of said information object.

45.   A method according to claim 44, wherein said changing of said information object is done in order to eliminate parts having policies that do not allow for said action to be executed while they are in the document.

46.   A method according to claim 44, wherein said changing of said information object is carried out in order to personalize said information object.

47.   A method according to claim 44, wherein said changing of said information object is carried out in order to customize said information object for a specific use.

48.   A method according to claim 44, wherein said changing of said

information object is done in a manner selected to achieve at least one of the following:

preserving the coherency of said information object; seamlessness; preserve the structure of said information object; preserving the linguistic coherency of said information object; preserving the formatting style of said information object; and preserve the pagination style of said information object.

49. A method according to claim 44, wherein said information objects comprise compound information objects and wherein said changing of said information object is made to constituent parts of a compound information object.

50. A method according to claim 38, wherein said storing comprises storage in at least one of the following:

a portable media device; a floppy disk; a hard drive; a portable hard drive; a flash card; a flash device; disk on key; magnetic tape; magnetic media; optic media; punched cards; a machine readable media; a CD; a DVD; a firewire device; a USB device; and a hand held computer.

51. A method according to claim 38, wherein said policy comprises distribution regulation, said distribution regulation being for regulating at least one of the following:

sending said information object via mail;

sending said information object via web mail;

uploading said information object to a web server;

uploading said information object to a FTP server;

sending said information object via a file transfer application;

sending said information object via an instant messaging application;

sending said information object via a file transfer protocol; and

sending said information object via an instant messaging protocol.

52. A method according to claim 38, wherein said policy is dependent on at least one of the following:

the domain of a respective information object; the identity of a system; the identity of a user; the identity level of a user authorizing an action; the identity of a user requesting an action; the identity of a user involved in an action; the identity of a user receiving an information object; the authentication

5   level of a system; the authentication level of a user; the authentication level of a user requesting an action; the authentication level of a user authorizing an action; the authentication level of a user involved in an action; the authentication level of a user receiving said information object; the authentication level of a user sending said information object; the format of an

10   information object instance; an interface being used; an application being used; encryption being used; digital rights management technology being used; detection of transformation, wherein said transformation is operable to reduce the ability to identify said transformed information object; information object integrity; regular usage pattern; regular distribution pattern; regular storage

15   pattern; information path; consistency of an action with usage pattern; the identity of a user overriding policy and authorizing the action in respect to said information object; the authentication level of a user overriding policy and authorizing the action in respect to said information object; the identity of a user sending information object; information property of said information object;

20   language of said information object; representation of said information object; operations done on of said information object; identity of users involved along the life cycle of said information object; application used on of said information object; transition channel of said information object; participant agents; virtual location of a computer; logical location of a computer; physical location of a

25   computer; type of a computer; type of a laptop computer; type of a desktop computer; type of a server computer; and owner identity.

53.   A method according to claim 38, further comprising enabling at least one user to override at least one of decisions contained within said policy.

30

54.   A method according to claim 1, wherein said deducing comprises utilizing conditional probabilities for at least one of the following:
    identification of information objects;
    classification of information objects; and

identification of a knowledge domain of information objects.

55.   A method according to claim 38, wherein at least part of said policy is stored in a database.

56.   A method according to claim 1, wherein said deducing further comprising utilizing keywords for at least one of the following:

identification of information objects; identification of elementary information units; classification of information objects; and identification of the domain of information objects.

57.   A method according to claim 56, wherein said keywords are stored in a database.

58.   A method according to claim 56, wherein said keywords are stored in at least one of the following forms:

hash value; raw string; and numeric representation.

59.   A method according to claim 38, wherein at least part of said policy is defined in terms of a logic expression.

60.   A method according to claim 59, wherein said expression is evaluated by lazy evaluation.

61.   A method according to claim 59, wherein at least some of the variables in said logic expression comprise of at least on of the following:

an external function; an external function based on group membership; and an external variable.

62. A method according to claim 38, wherein at least part of said policy is defined in terms of any one of a group comprising rules, imposed restrictions, granted privileges, reaction to one or more given events, group operations, a

property of said information object, a property of a user, a property of a computer, a property of an entity, and a hierarchy of calculations.

63.   A method according to claim 38, wherein at least part of said policy is defined in terms of a role, wherein said role consists of a property of at least one of a user and a system and wherein said role further comprises at least one authorization.

64.   A method according to claim 38, wherein at least part of said policy is defined in terms of at least one of the following languages:
a scripting language; an ordered calculation language; a programming language; an interpreted language; and a functional language.

65.   A method according to claim 64, wherein said at least one of said following languages comprises instructions for the operation of an ordered calculation resulting in at least one of the following:
policy; instruction to perform an action; restriction; and allowance.

66.   A method according to claim 38, wherein said information object is a compound information object comprising constituent simple information objects, and a respective policy assigned to said information object comprises different policies for at least some of said constituent information objects.

67.   A method according to claim 1, wherein at least one user is defined in an owner definition as an owner of said information object.

68.   A method according to claim 67, wherein said owner definition is stored in a database.

69.   A method according to claim 1, wherein said deducing further comprises utilizing organizational structure information.

70.   A method according to claim 69, wherein said organizational structure information comprise at least one of the following:

user superiority; working groups; organizational hierarchy; departmental separation; and membership in working groups.

71.   A method according to claim 38, comprising using organizational structure information in order to assign a respective policy object.

72.   A method according to claim 69, wherein at least part of said organizational structure information is stored in a database.

73.   A method according to claim 69, wherein at least part of said organizational structure information is used for information object classification.

74.   A method according to claim 69, wherein at least part of said organizational structure information is imported from at least one of the following:

organizational data system; data management system; organizational data management system;

knowledge management system; user directory; LDAP server; document; and an organizational chart.

75.   A method according to claim 1, further comprising making use of at least one user interface operable to assist in at least one of the following:

classification; policy definition; template definition; approving and revising automatic template definition; importing organizational structure information; revising organizational structure information; produce reports; overriding policy decisions; and providing authorizations.

76.   A method according to claim 38, comprising defining an information class as a group consisting of at least two information objects, said

defining further comprising associating with said information class a corresponding class policy being a policy shared by said information objects.

77.   A method according to claim 76, wherein said information class policy comprises at least a part of respective policies of said information objects within said class.

78.   A method according to claim 1, further comprising using template information objects to represent commonly repeated information, such that a template information object together with a difference information object representing instance specific information are together formable to produce a compound information object in which common and specific information are respectively identifiable.

79.   A method according to claim 78, comprising using said template information object in identifying any of unknown information object comprising information corresponding to said template information object.

80.   A method according to claim 78, wherein said template information object is a compound information object, wherein said template information object comprises at least one placeholder, and wherein said method comprises replacing said placeholder by at least part of said difference information object when said difference information object and a respective template information object are combined.

81.   A method according to claim 80, wherein at least one of said placeholders is a specialized placeholder, said specialized placeholder comprising specialization information to identify a respective specialization of said specialized placeholder.

82.   A method according to claim 80, wherein at least one of said placeholders is a specialized placeholder, said specialized placeholder comprising a restriction about information objects permitted for replacing said specialized placeholder.

83.   A method according to claim 78, wherein said template information object comprises at least one of a group comprising: a disclaimer; a form; a header; a footer; a contract; and an invoice.

84.   A method according to claim 81, wherein said specialized placeholder comprises a restriction about information objects permitted for replacing said specialized placeholder, and wherein said restriction comprises a rule for excluding at least one of the following:

an object comprising numeric information;

an object comprising a word;

an object comprising a character;

an object comprising a digit;

an object comprising a sentence; and

an object comprising a simple information object.

85.   A method according to claim 78, comprising defining a template information object and wherein said defining comprises automatically identifying a template information object candidate.

86.   A method according to claim 85, wherein said automatically identifying a template information object candidate comprises identification of shared elementary information units of at least two information objects.

87.   A method according to claim 85, wherein said step of automatically identifying a template information object candidate comprises identification of substantially similar information objects.

88.   A method according to claim 85, wherein said step of automatically identifying a template information object candidate comprises the use of at least one of text parsing; and text matching.

89.   A method according to claim 78, comprising deriving at least a part of a respective information object policy associated with a template

instance information object from an information object policy of the respective originating template information object.

90. A method according to claim 38, wherein at least a part of an information object policy of a respective information object is derived from a default information object policy when said part of said information object policy of said information object is not explicitly defined.

91. A method according to claim 5, comprising applying preprocessing to said elementary information units before assigning identifiers thereto.

92. A method according to claim 91, wherein said preprocessing is done in order to enhance at least one of efficiency and robustness.

93. A method according to claim 91, wherein said preprocessing comprises at least one of canonization; removal of common words; removal of words not having a substantial effect on the meaning of the text; removal of punctuation; correction of spelling; canonization of spelling; scene detection; canonizing size; canonizing orientation; canonizing color; removing color; reducing noise; enhancing area separation; enhancing borders; enhancing lines; sharpening; blurring; removal of elementary information units substantially similar to neighboring elementary information units; canonization of grammar; and transformation to a phonetic representation.

94. A method according to claim 91, comprising carrying out said preprocessing so as to ensure that any area of a given size in said information object contains at least a predetermined number of said elementary information units having an assigned elementary information unit identifier.

95. A method according to claim 94, wherein said given size is dependent on properties of said information object.

96. A method according to claim 95, wherein said properties of said information object comprise at least one of a group comprising: importance; size; confidentiality level; and format.

97. A method according to claim 94, wherein said predetermined number is dependent on properties of said information object.

98. A method according to claim 97, wherein said properties of said information object comprise at least one of a group comprising: importance; size; confidentiality level; and format.

99. A method according to claim 1, further comprising a stage of detection of information objects having undergone transformations.

100. A method according to claim 99, wherein said stage of detection is aimed for detection of transformations intended to reduce the ability to identify said information object.

101. A method according to claim 99, wherein said detection of information objects that have undergone transformation comprises detection of at least one of a group comprising:

transformation artifacts; spelling mistakes; wrong grammar; wrong punctuation; wrong capitalization; missing punctuation; missing capitalization; irregular word distribution; lack of common words; predominance of unknown words; inconsistent headers; headers inconsistent with file type; headers inconsistent with file content; file type inconsistent with file content; irregular distribution of characters; irregular distribution of words; irregular distribution of character sequences; irregular distribution of word sequences; irregular length of words; irregular length of sentences; irregular distribution of length of words; irregular distribution of length of sentences; irregular file format; irregular file encoding; unknown file format; unknown file encoding; mix of non-alphabetic characters; unopenable file; action time; information object

creation time; information object update time; encryption; and an unexpectedly high level of entropy.

102.   A method according to claim 5, comprising formulating respective assigned elementary information unit identifier to be resilient to small errors.

103.   A method according to claim 5, wherein said assigning of elementary information unit identifier utilizes image matching.

104.   A method according to claim 5, wherein said assigning of elementary information unit identifier comprises a mapping to a Euclidian space.

105.   A method according to claim 104, wherein said mapping to a Euclidian space comprises approximating a pairwise difference between elementary information units.

106.   A method according to claim 105, wherein said approximating is such that a difference between two elementary information units approximates said pairwise difference between said two elementary information units.

107.   A method according to claim 105, wherein said approximation of said pairwise difference between elementary information units comprises an approximation of at least one of the following:
semantic difference; distance measured by image matching; phonetic difference; and spelling difference.

108.   A method according to claim 1, wherein said information object is a knowledge object.

109.   A method according to claim 1, wherein said elementary information unit is an elementary fact.

110. A method according to claim 109, wherein said elementary fact comprises at least one of the following:

sentence; database entry; representation independent description of knowledge; modular description of knowledge; and abstract description of knowledge.

111. A method according to claim 76, wherein said information class is a knowledge class.

112. A method according to claim 1, further comprising a stage of discerning lifecycle information about a respective information object.

113. A method according to claim 112, wherein said discerning of information about the lifecycle of said information object comprises utilizing information about sharing of at least one elementary information unit in said information object, wherein said elementary information unit is shared with at least one additional information object.

114. A method according to claim 112, wherein said discerning of information about the lifecycle of said information object is based on at least one of a group comprising: file system date information; information about editing of said information object; and information about registration of said information object.

115. A method according to claim 112, comprising utilizing said information about the lifecycle of said information object for the creation of a lifecycle graph.

116. A method according to claim 112, comprising utilizing said information about the lifecycle of said information object to define at least part of the policy of said information object said utilizing comprising identifying at least one other information object along said information object's lifecycle and examining a policy associated therewith.

117. A method according to claim 5, wherein said assigning of said elementary information unit identifier is carried out a plurality of times, each time utilizing a different method for assigning of an elementary information unit identifier.

5

118. A method according to claim 117, wherein said assigning of elementary information unit identifier several times comprises storing of elementary information unit identifier assigned utilizing different methods are stored separately.

10

119. A method according to claim 117, wherein said assigning of said elementary information unit identifier several times comprises storing of elementary information unit identifier assigned utilizing different methods can be distinguished according to said method utilized to assign them.

15

120. A method according to claim 117, wherein said different methods are selected such as to optimize between at least any two of the following:

storage space; search speed; capability to detect transformation;
20 capability to detect a specific transformation; resilience to transformation; resolution of identification from among similar information objects; resolution of identification of boundaries within compound information objects; resilience to a specific transformation; and resilience to transformation.

25 121. A method according to claim 5, wherein said assigning of a respective elementary information unit identifier comprises utilizing a method having at least one of the following characteristics:

order sensitive to data in the elementary information unit; order insensitive in the elementary information unit; utilizing changing definitions of
30 the elementary information unit such that said assigning of said elementary information unit identifier is carried out a plurality of times using a plurality of definitions; utilizing an exchangeable method of preprocessing, such that said assigning of said elementary information unit identifier is carried out several times; being omission resilient; being insertion resilient; being replacement

resilient; being dictionary based; being distribution based; being locality based; being histogram based; and being n-gram based.

122. A method according to claim 38, wherein an information object policy comprises at least some information about one or more methods utilized for assigning of an elementary information unit identifier to a respective information object.

123. A method according to claim 117, wherein said assigning utilizing different methods comprises utilizing said different methods sequentially until a predetermined stop condition is reached.

124. A method according to claim 5, wherein said information object comprises spreadsheet data, and wherein said assigning of said elementary information unit identifier assigned to said information object comprises utilizing a method comprising at least one of the following characteristics:
invariance to linear transformation; invariance to reordering; invariance to permutation; resilience to linear transformation; resilience to reordering; resilience to permutation; resilience to minor changes; resilience to cuts; utilizing of statistic moment; utilizing of statistic moment for a table; utilizing statistic moment for a row; utilizing statistic moment for a column; and utilizing a mathematical descriptor of the information object data.

125. A method according to claim 5, comprising utilizing said elementary information unit identifiers for said information object identification using a technique having at least one of the following characteristics: omission resilience; insertion resilience; replacement resilience; being dictionary based; being distribution based; being locality based; being based on the size of elementary information units; being based on the size of information objects; resilience to linear transformation; resilience to reordering; resilience to permutation; resilience to minor changes; resilience to cuts; being histogram based; and being n-gram based.

126.   A method according to claim 1, further comprising utilizing a client.

127.   A method according to claim 126, wherein said client comprises at least one of the following:

end point software; end point hardware; tamper resistant software; tamper resistant hardware;

client side software; and client side hardware.

128.   A method according to claim 126, comprising utilizing said client for at least one of the following:

monitoring of client side storage; monitoring of client side access; monitoring of client side usage;

monitoring of client side distribution;

monitoring of copying of information object excerpts;

monitoring of clipboard;

monitoring of at least one application;

monitoring of at least one interface;

control of at least one application;

control of at least one interface;

control of clipboard;

control of copying of information object excerpts;

control of client side storage;

control of client side access;

control of client side usage; and

control of client side distribution.

129.   A method according to claim 1, comprising utilizing comparing of at least two information objects to calculate pairwise similarity between objects.

130.   A method according to claim 129, comprising utilizing said pairwise similarity to map said information objects to a space.

131. A method according to claim 130, wherein said space is an Euclidian space, and wherein the closeness between any two objects within said Euclidian space is approximately proportional to said pairwise similarity between said information objects.

5

132. A method according to claim 130, wherein said space is a weighted graph, and wherein the weight of an edge between any two objects within said graph space is approximately proportional to said pairwise similarity between said information objects.

10

133. A method according to claim 130, wherein said space is a graph, and wherein the existence of an edge between any two objects within said graph space is dependent on said pairwise similarity between said information objects.

15

134. A method according to claim 130, wherein said space is utilized to identify at least one similarity information class, wherein said information class consists of at least two information objects, wherein said information class policy is a policy shared by the information class, and wherein said similarity information class is bounded within said space.

20

135. A method according to claim 130, comprising utilizing said space to identify at last one information object substantially similar to an unidentified information object.

25

136. A method according to claim 130, comprising using said space to identify at least one other information object substantially similar to an information object for which policy is not known, thereby to obtain a policy associated with said other information object to use as basis for a policy for said information object.

30

137. A method according to claim 1, comprising storing information about said information object in a database.

138.   A method according to claim 1, further comprising extracting a descriptor of said information object, based on statistical analysis of said information object.

5   139.   A method according to claim 1, comprising storing the order of said elementary information units within said information object in a database.

140.   A method according to claim 139, comprising using said order for identification of said information object.

10

141.   A method according to claim 1, further comprising interfacing at least one of an information management system; and a document management system.

15   142.   A method according to claim 1, further comprising tracking at least one of the following:

usage patterns; storage patterns; and distribution patterns.

143.   A method according to claim 142, wherein said tracking is
20   carried out to infer information about at least one of the following:

normal usage patterns; normal storage patterns; normal distribution patterns; irregular usage patterns; irregular storage patterns; and irregular distribution patterns.

25   144.   A method according to claim 143, wherein said inferred information is used to define at least part of a policy.

145.   A method according to claim 143, comprising using said inferred information for information object classification.

30

146.   A method according to claim 1, further comprising logging.

147.   A method according to claim 146, wherein said logging comprising logging of at least one of the following:

actions; events; and information objects identification.

148. A method according to claim 146, wherein at least part of said logging is controlled by a policy.

149. A method according to claim 146, wherein at least part of said logging is stored in a database.

150. A method according to claim 146, comprising utilizing said logging to augment lifecycle information for said information object.

151. A method according to claim 1, further comprising assessing the integrity of at least one information object, wherein said integrity assessment consists of comparing said information object with a version of said information object for which integrity is assured.

152. A method according to claim 151, further comprising issuing a certificate of said integrity for at least one information object.

153. A method according to claim 152, wherein said certificate is a cryptographic certificate.

154. A method according to claim 151, further comprising replacing said information object with said version of said information object for which said integrity is assured.

155. A method according to claim 151, comprising identifying when said integrity of said information object is not satisfactory, and in such a case not allowing distribution of said information object.

156. A method according to claim 151, comprising identifying when said integrity of said information object is not satisfactory, and in such a case not allowing storage of said information object..

157. A method according to claim 151, comprising identifying when said integrity of said information object is not satisfactory, and in such a case not allowing usage of said information object.

5

158. A method according to claim 1, further comprising defining at least one constituent information object to be an ignored information object, and wherein, whenever said to be ignored information object is an element of a compound information, ignoring said object in identification of said compound information object.

10

159. A method according to claim 38, further comprising changing access control information in accordance with said policy.

160. A method according to claim 1, further comprising not allowing

15   usage of respective ones of said information objects outside an organization.

161. A method according to claim 1, further comprising not allowing storage of respective ones of said information object outside an organization.

20

162. A method according to claim 1, further comprising not allowing distribution of respective ones of said information object outside an organization.

163. A method according to claim 40, wherein said policy comprises

25   at least one mandatory lifecycle.

164. A method according to claim 163, wherein said action is dependent on the matching of said mandatory lifecycle with a lifecycle of a respective event.

30

165. A method according to claim 163, wherein said mandatory lifecycle comprises at least one mandatory recipient of said information object; and an order of events concerning said information object.

166. A method according to claim 44, wherein said inserting an additional part to said information object comprises inserting at least one of the following: a header; a footer; and a disclaimer.

5   167. A method according to claim 38 or claim 52, comprising defining areas and wherein said policy is dependent on whether an action is taken inside a user-defined area.

168. A method according to claim 38 or claim 52, comprising
10   defining areas and wherein said policy is dependent on whether an event occurs inside a user-defined area.

169. A method according to claim 1 or claim 5, comprising using said deducing to locate at least one information object with similar content to a given
15   information object.

170. A method according to claim 38, comprising attaching a respective policy to information objects according to their logical location within an information storage medium.
20

171. A method according to claim 170, further comprising utilizing a crawler for automatic location of information objects.

172. A method according to claim 171, wherein said information
25   storage medium is a file system.

173. A method according to claim 169, wherein said locating is done in an information storage medium.
30

174. A method according to claim 173, further comprising utilizing a crawler for automatic location of information objects within said information storage medium.

35   175. A method according to claim 173, wherein said information storage medium comprises at least one file system.

176.   A method for information identification comprising:

Finding elementary information units within said information object; and

5   Deducing information about the identity of said information object from identification of said elementary information units found within said information object.

177.   A method according to claim 176, wherein said information

10   objects comprise at least one simple information object, said simple information object comprising one of the following:

an elementary information unit;

a set of elementary information units; and

an ordered set of elementary information units.

15

178.   A method according to claim 176, wherein said elementary information units comprise at least one of the following:

a sentence; a sequences of words; a word; a sequence of characters; a character; a sequence of numbers; a number; a sequence of digits; a digit; a

20   vector; a curve; a pixel; a block of pixels; an audio frame; a musical note; a musical bar; a visual object; a sequence of video frames; a sequence of musical notes; a sequence of musical bars; and a video frame.

179.   A method according to claim 176, further comprising assigning

25   elementary information units identifiers to elementary information units after identification.

180.   A method according to claim 179, wherein said elementary

30   information unit identifiers are utilized in said deducing.

181.   A method according to claim 176, wherein said information object identification is carried out on an instance of said information object, said information object instance being said information object in a specific format.

182.   A method according to claim 181, wherein said format comprise at least one of the following:

jpeg image; gif image; Word document format; Lotus notes format; mpeg format; text format; rich text format; Unicode text format; multi byte text encoding format; formatted text format; ASCII text format; HTML; XML; PDF; postscript; MS-Excel spreadsheet; MS-Excel drawing; MS-Visio drawing; Photoshop drawing; AutoCAD drawing format; and CAD drawing format.

183.   A method according to claim 179, wherein said elementary information unit identifiers are determined by the content of said elementary information units which they are assigned to.

184.   A method according to claim 183, wherein said elementary information unit identifiers are solely determined by said content.

185.   A method according to claim 179, wherein said elementary information units identifiers are at least partly determined by locations within an information object of respective elementary information units to which they are assigned.

186.   A method according to claim 179, wherein said elementary information units identifiers are at least partly determined by the content of an elementary information unit in proximity to said elementary information units to which they are assigned.

187.   A method according to claim 179, comprising storing said elementary information units identifiers in a database.

79

188.   A method according to claim 187, further comprising using said elementary information units identifiers stored in said database for identifying at least one further, unidentified, information object.

5   189.   A method according to claim 187, further comprising using said elementary information units identifiers stored in said database for comparing information objects.

190.   A method according to claim 179, comprising storing only some

10   of said elementary information units identifiers in a database.

191.   A method according to claim 190, wherein said storing of only some of said elementary information units identifiers in a database is to achieve at least one of the following:

15   reduce storage cost;

increase efficiency of assigning of said elementary information units identifiers to said elementary information units by only performing said assignment for elementary information units identifiers that are stored in said database; and

20   increase the efficiency of searching for said elementary information units identifiers in said database.

192.   A method according to claim 190, wherein said storage of only some of said elementary information units identifiers in a database is done in a

25   manner that ensures that any area of a given size in said information object contains a predetermined minimum number of said stored elementary information units.

193.   A method according to claim 192, wherein said given size is

30   dependent on properties of a respective information object.

194.   A method according to claim 193, wherein said properties of said information object comprise at least one of the following:

importance; size; confidentiality level; and format.

80

195. A method according to claim 192, wherein said minimum number is dependent on properties of said information object.

5   196. A method according to claim 195, wherein said properties of said information object comprise at least one of the following:

importance; size; confidentiality level; and format.

197. A method according to claim 177, wherein said information

10   objects comprise at least one compound information object, said compound information object comprising at least one of the following:

a simple information object; a compound information object; an ordered set of compound information objects; an ordered set of simple information objects; and an ordered set of compound and simple information objects.

15

198. A method according to claim 176, wherein said information comprises at least one of the following:

numeric data; spreadsheet data; numeric spreadsheet data; textual spreadsheet data; word processor data; textual data; hyper text data; audio data;

20   visual data; multimedia data; binary data; raw data; database data; video data; drawing data; chart data; picture data; and image data.

199. A method according to claim 176, comprising using said deducing to attach to said information object an information object policy, said

25   policy comprising at least one of the following:

an allowed distribution of said information object;

a restriction on distribution of said information object;

an allowed storage of said information object;

a restriction on storage of said information object;

30   an action to be taken as a reaction to an event;

an allowed usage of said information object; and

a restriction on usage of said information object.

200.   A method according to claim 199, wherein said information object policy comprises at least one action to be taken as a reaction to an event, and wherein said action comprises at least one of the following:

preventing distribution of said information object; preventing storage of said information object; preventing usage of said information object; reporting distribution of said information object; reporting storage of said information object; reporting usage of said information object; reporting; alerting about distribution of said information object; alerting storage of said information object; alerting usage of said information object; alerting; logging distribution of said information object; logging storage of said information object; logging usage of said information object; logging; notifying about distribution of said information object; notifying about storage of said information object; notifying about usage of said information object; notifying; notifying to an administrator; notifying to a manager; notifying to a recipient; notifying to a sender; notifying to an owner of said information object; quarantine; alerting an administrator; alerting a manager; alerting a recipient; alerting a sender; alerting an owner of said information object; reporting to an administrator; reporting to a manager; reporting to a recipient; reporting to a sender; reporting to an owner of said information object; encrypting said information object; changing said information object; replacing said information object; and utilizing digital rights management technology on said information object.

201.   A method according to claim 199, wherein said information object policy comprises at least one action to be taken as a reaction to an event, and wherein said event comprises at least one of the following:

attempted distribution of said information object; attempted storage of said information object;
attempted usage of said information object; distribution of said information object; storage of said information object; and usage of said information object.

202.   A method according to claim 199, wherein said information object usage comprises at least one of the following:

82

copying an excerpt; editing; copying to clipboard; copying an excerpt to clipboard; changing format; changing encoding; encryption; decryption; changing digital management; opening by an application; and printing.

203.   A method according to claim 199, wherein said information object policy comprises placing a substantially imperceptible marking in said information object, said marking comprising information content, and said method comprising placing said marking, when indicated by said policy, before allowing at least one of the following:

storage of said information object; usage of said information object; and distribution of said information object.

204.   A method according to claim 203, wherein said information content for storage in said marking comprises at least one of the following:

the identity of said information object;

the identity of a user performing the action in respect to said information object;

the identity of a user authorizing the action in respect to said information object;

the identity of a user overriding policy and approving the action in respect to said information object; and

the identity of a user requesting the action in respect to said information object.

205.   A method according to claim 199, wherein said information object policy further comprises changing said information object by at least one of the following:

deleting part of said information object; replacing part of said information object; and inserting an additional part to said information object

before allowing at least one of the following actions:

storage of said information object; usage of said information object; and distribution of said information object.

206. A method according to claim 205, wherein said changing of said information object is done in order to eliminate parts having policies that do not allow for said action to be executed while they are in the document.

5   207. A method according to claim 205, wherein said changing of said information object is carried out in order to personalize said information object.

208. A method according to claim 205, wherein said changing of said information object is carried out in order to customize said information object

10   for a specific use.

209. A method according to claim 205, wherein said changing of said information object is done in a manner selected to achieve at least one of the following:

15   preserving the coherency of said information object; seamlessness; preserve the structure of said information object; preserving the linguistic coherency of said information object; preserving the formatting style of said information object; and preserve the pagination style of said information object.

20   210. A method according to claim 205, wherein said information objects comprise compound information objects and wherein said changing of said information object is made to constituent parts of a compound information object.

25   211. A method according to claim 199, wherein said storing comprises storage in at least one of the following:
          a portable media device; a floppy disk; a hard drive; a portable hard drive; a flash card; a flash device; disk on key; magnetic tape; magnetic media; optic media; punched cards; a machine readable media; a CD; a DVD; a

30   firewire device; a USB device; and a hand held computer.

212. A method according to claim 199, wherein said policy comprises distribution regulation, said distribution regulation being for regulating at least one of the following:

sending said information object via mail;

sending said information object via web mail;

uploading said information object to a web server;

uploading said information object to a FTP server;

sending said information object via a file transfer application;

sending said information object via an instant messaging application;

sending said information object via a file transfer protocol; and

sending said information object via an instant messaging protocol.

213. A method according to claim 199, wherein said policy is dependent on at least one of the following:

the domain of a respective information object; the identity of a system; the identity of a user; the identity level of a user authorizing an action; the identity of a user requesting an action; the identity of a user involved in an action; the identity of a user receiving an information object; the authentication level of a system; the authentication level of a user; the authentication level of a user requesting an action; the authentication level of a user authorizing an action; the authentication level of a user involved in an action; the authentication level of a user receiving said information object; the authentication level of a user sending said information object; the format of an information object instance; an interface being used; an application being used; encryption being used; digital rights management technology being used; detection of transformation, wherein said transformation is operable to reduce the ability to identify said transformed information object; information object integrity; regular usage pattern; regular distribution pattern; regular storage pattern; information path; consistency of an action with usage pattern; the identity of a user overriding policy and authorizing the action in respect to said information object; the authentication level of a user overriding policy and authorizing the action in respect to said information object; the identity of a user sending information object; information property of said information object; language of said information object; representation of said information object;

operations done on of said information object; identity of users involved along the life cycle of said information object; application used on of said information object; transition channel of said information object; participant agents; virtual location of a computer; logical location of a computer; physical location of a

5   computer; type of a computer; type of a laptop computer; type of a desktop computer; type of a server computer; and owner identity.

214. A method according to claim 199, further comprising enabling at least one user to override at least one of decisions contained within said policy.

10

215. A method according to claim 176, wherein said deducing comprises utilizing conditional probabilities for at least one of the following:
        identification of information objects;
        classification of information objects; and
15   identification of a knowledge domain of information objects.

216. A method according to claim 199, wherein at least part of said policy is stored in a database.

20   217. A method according to claim 176, wherein said deducing further comprising utilizing keywords for at least one of the following:
        identification of information objects; identification of elementary information units; classification of information objects; and identification of the domain of information objects.

25

218. A method according to claim 217, wherein said keywords are stored in a database.

219. A method according to claim 217, wherein said keywords are
30   stored in at least one of the following forms:
        hash value; raw string; and numeric representation.

220. A method according to claim 199, wherein at least part of said policy is defined in terms of a logic expression.

221. A method according to claim 220, wherein said expression is evaluated by lazy evaluation.

5   222. A method according to claim 220, wherein at least some of the variables in said logic expression comprise of at least on of the following:

an external function; an external function based on group membership; and an external variable.

10   223. A method according to claim 199, wherein at least part of said policy is defined in terms of any one of a group comprising rules, imposed restrictions, granted privileges, reaction to one or more given events, group operations, a property of said information object, a property of a user, a property of a computer, a property of an entity, and a hierarchy of calculations.

15

224. A method according to claim 199, wherein at least part of said policy is defined in terms of a role, wherein said role consists of a property of at least one of a user and a system and wherein said role further comprises at least one authorization.

20

225. A method according to claim 199, wherein at least part of said policy is defined in terms of at least one of the following languages:

a scripting language; an ordered calculation language; a programming language; an interpreted language; and a functional language.

25

226. A method according to claim 225, wherein said at least one of said following languages comprises instructions for the operation of an ordered calculation resulting in at least one of the following:

policy; instruction to perform an action; restriction; and allowance.

30

227. A method according to claim 199, wherein said information object is a compound information object comprising constituent simple

information objects, and a respective policy assigned to said information object comprises different policies for at least some of said constituent information objects.

228. A method according to claim 176, wherein at least one user is defined in an owner definition as an owner of said information object.

229. A method according to claim 228, wherein said owner definition is stored in a database.

230. A method according to claim 176, wherein said deducing further comprises utilizing organizational structure information.

231. A method according to claim 230, wherein said organizational structure information comprise at least one of the following:
user superiority; working groups; organizational hierarchy; departmental separation; and membership in working groups.

232. A method according to claim 199, comprising using organizational structure information in order to assign a respective policy object.

233. A method according to claim 230, wherein at least part of said organizational structure information is stored in a database.

234. A method according to claim 230, wherein at least part of said organizational structure information is used for information object classification.

235. A method according to claim 230, wherein at least part of said organizational structure information is imported from at least one of the following:
organizational data system; data management system; organizational data management system;

knowledge management system; user directory; LDAP server; document; and an organizational chart.

236.   A method according to claim 176, further comprising making use of at least one user interface operable to assist in at least one of the . following:

classification; policy definition; template definition; approving and revising automatic template definition; importing organizational structure information; revising organizational structure information; produce reports; overriding policy decisions; and providing authorizations.

237.   A method according to claim 199, comprising defining an information class as a group consisting of at least two information objects, said defining further comprising associating with said information class a corresponding class policy being a policy shared by said information objects.

238.   A method according to claim 237, wherein said information class policy comprises at least a part of respective policies of said information objects within said class.

239.   A method according to claim 176, further comprising using template information objects to represent commonly repeated information, such that a template information object together with a difference information object representing instance specific information are together formable to produce a compound information object in which common and specific information are respectively identifiable.

240.   A method according to claim 239, comprising using said template information object in identifying any of unknown information object comprising information corresponding to said template information object.

241.   A method according to claim 239, wherein said template information object is a compound information object, wherein said template information object comprises at least one placeholder, and wherein said method

comprises replacing said placeholder by at least part of said difference information object when said difference information object and a respective template information object are combined.

5   242.   A method according to claim 241, wherein at least one of said placeholders is a specialized placeholder, said specialized placeholder comprising specialization information to identify a respective specialization of said specialized placeholder.

10   243.   A method according to claim 241, wherein at least one of said placeholders is a specialized placeholder, said specialized placeholder comprising a restriction about information objects permitted for replacing said specialized placeholder.

15   244.   A method according to claim 239, wherein said template information object comprises at least one of a group comprising: a disclaimer; a form; a header; a footer; a contract; and an invoice.

        245.   A method according to claim 242, wherein said specialized
20   placeholder comprises a restriction about information objects permitted for replacing said specialized placeholder, and wherein said restriction comprises a rule for excluding at least one of the following:
        an object comprising numeric information;
        an object comprising a word;
25   an object comprising a character;
        an object comprising a digit;
        an object comprising a sentence; and
        an object comprising a simple information object.

30   246.   A method according to claim 239, comprising defining a template information object and wherein said defining comprises automatically identifying a template information object candidate.

247. A method according to claim 246, wherein said automatically identifying a template information object candidate comprises identification of shared elementary information units of at least two information objects.

5   248. A method according to claim 246, wherein said step of automatically identifying a template information object candidate comprises identification of substantially similar information objects.

249. A method according to claim 246, wherein said step of
10   automatically identifying a template information object candidate comprises the use of at least one of text parsing; and text matching.

250. A method according to claim 239, comprising deriving at least a part of a respective information object policy associated with a template
15   instance information object from an information object policy of the respective originating template information object.

251. A method according to claim 199, wherein at least a part of an information object policy of a respective information object is derived from a
20   default information object policy when said part of said information object policy of said information object is not explicitly defined.

252. A method according to claim 179, comprising applying preprocessing to said elementary information units before assigning identifiers
25   thereto.

253. A method according to claim 252, wherein said preprocessing is done in order to enhance at least one of efficiency and robustness.

30   254. A method according to claim 252, wherein said preprocessing comprises at least one of canonization; removal of common words; removal of words not having a substantial effect on the meaning of the text; removal of punctuation; correction of spelling; canonization of spelling; scene detection; canonizing size; canonizing orientation; canonizing color; removing color;

reducing noise; enhancing area separation; enhancing borders; enhancing lines; sharpening; blurring; removal of elementary information units substantially similar to neighboring elementary information units; canonization of grammar; and transformation to a phonetic representation.

255.   A method according to claim 252, comprising carrying out said preprocessing so as to ensure that any area of a given size in said information object contains at least a predetermined number of said elementary information units having an assigned elementary information unit identifier.

256.   A method according to claim 255, wherein said given size is dependent on properties of said information object.

257.   A method according to claim 256, wherein said properties of said information object comprise at least one of a group comprising: importance; size; confidentiality level; and format.

258.   A method according to claim 255, wherein said predetermined number is dependent on properties of said information object.

259.   A method according to claim 258, wherein said properties of said information object comprise at least one of a group comprising: importance; size; confidentiality level; and format.

260.   A method according to claim 176, further comprising a stage of detection of information objects having undergone transformations.

261.   A method according to claim 260, wherein said stage of detection is aimed for detection of transformations intended to reduce the ability to identify said information object.

262. A method according to claim 260, wherein said detection of information objects that have undergone transformation comprises detection of at least one of a group comprising:

transformation artifacts; spelling mistakes; wrong grammar; wrong punctuation; wrong capitalization; missing punctuation; missing capitalization; irregular word distribution; lack of common words; predominance of unknown words; inconsistent headers; headers inconsistent with file type; headers inconsistent with file content; file type inconsistent with file content; irregular distribution of characters; irregular distribution of words; irregular distribution of character sequences; irregular distribution of word sequences; irregular length of words; irregular length of sentences; irregular distribution of length of words; irregular distribution of length of sentences; irregular file format; irregular file encoding; unknown file format; unknown file encoding; mix of non-alphabetic characters; unopenable file; action time; information object creation time; information object update time; encryption; and an unexpectedly high level of entropy.

263. A method according to claim 179, comprising formulating respective assigned elementary information unit identifier to be resilient to small errors.

264. A method according to claim 179, wherein said assigning of elementary information unit identifier utilizes image matching.

265. A method according to claim 179, wherein said assigning of elementary information unit identifier comprises a mapping to a Euclidian space.

266. A method according to claim 265, wherein said mapping to a Euclidian space comprises approximating a pairwise difference between elementary information units.

267.   A method according to claim 266, wherein said approximating is such that a difference between two elementary information units approximates said pairwise difference between said two elementary information units.

5   268.   A method according to claim 266, wherein said approximation of said pairwise difference between elementary information units comprises an approximation of at least one of the following:

semantic difference; distance measured by image matching; phonetic difference; and spelling difference.

10

269.   A method according to claim 176, wherein said information object is a knowledge object.

270.   A method according to claim 176, wherein said elementary

15   information unit is an elementary fact.

271.   A method according to claim 270, wherein said elementary fact comprises at least one of the following:

sentence; database entry; representation independent description of

20   knowledge; modular description of knowledge; and abstract description of knowledge.

272.   A method according to claim 237, wherein said information class is a knowledge class.

25

273.   A method according to claim 176, further comprising a stage of discerning lifecycle information about a respective information object.

274.   A method according to claim 273, wherein said discerning of

30   information about the lifecycle of said information object comprises utilizing information about sharing of at least one elementary information unit in said information object, wherein said elementary information unit is shared with at least one additional information object.

275. A method according to claim 273, wherein said discerning of information about the lifecycle of said information object is based on at least one of a group comprising: file system date information; information about editing of said information object; and information about registration of said information object.

276. A method according to claim 273, comprising utilizing said information about the lifecycle of said information object for the creation of a lifecycle graph.

277. A method according to claim 273, comprising utilizing said information about the lifecycle of said information object to define at least part of the policy of said information object said utilizing comprising identifying at least one other information object along said information object's lifecycle and examining a policy associated therewith.

278. A method according to claim 179, wherein said assigning of said elementary information unit identifier is carried out a plurality of times, each time utilizing a different method for assigning of an elementary information unit identifier.

279. A method according to claim 278, wherein said assigning of elementary information unit identifier several times comprises storing of elementary information unit identifier assigned utilizing different methods are stored separately.

280. A method according to claim 278, wherein said assigning of said elementary information unit identifier several times comprises storing of elementary information unit identifier assigned utilizing different methods can be distinguished according to said method utilized to assign them.

281. A method according to claim 278, wherein said different methods are selected such as to optimize between at least any two of the following:

storage space; search speed; capability to detect transformation; capability to detect a specific transformation; resilience to transformation; resolution of identification from among similar information objects; resolution of identification of boundaries within compound information objects; resilience

5   to a specific transformation; and resilience to transformation.

282.   A method according to claim 179, wherein said assigning of a respective elementary information unit identifier comprises utilizing a method having at least one of the following characteristics:

10   order sensitive to data in the elementary information unit; order insensitive in the elementary information unit; utilizing changing definitions of the elementary information unit such that said assigning of said elementary information unit identifier is carried out a plurality of times using a plurality of definitions; utilizing an exchangeable method of preprocessing, such that said

15   assigning of said elementary information unit identifier is carried out several times; being omission resilient; being insertion resilient; being replacement resilient; being dictionary based; being distribution based; being locality based; being histogram based; and being n-gram based.

20   283.   A method according to claim 199, wherein an information object policy comprises at least some information about one or more methods utilized for assigning of an elementary information unit identifier to a respective information object.

25.   284.   A method according to claim 278, wherein said assigning utilizing different methods comprises utilizing said different methods sequentially until a predetermined stop condition is reached.

285.   A method according to claim 179, wherein said information

30   object comprises spreadsheet data, and wherein said assigning of said elementary information unit identifier assigned to said information object comprises utilizing a method comprising at least one of the following characteristics:

96

invariance to linear transformation; invariance to reordering; invariance to permutation; resilience to linear transformation; resilience to reordering; resilience to permutation; resilience to minor changes; resilience to cuts; utilizing of statistic moment; utilizing of statistic moment for a table; utilizing

5   statistic moment for a row; utilizing statistic moment for a column; and utilizing a mathematical descriptor of the information object data.

286.   A method according to claim 179, comprising utilizing said elementary information unit identifiers for said information object identification

10   using a technique having at least one of the following characteristics: omission resilience; insertion resilience; replacement resilience; being dictionary based; being distribution based; being locality based; being based on the size of elementary information units; being based on the size of information objects; resilience to linear transformation; resilience to reordering; resilience to

15   permutation; resilience to minor changes; resilience to cuts; being histogram based; and being n-gram based.

287.   A method according to claim 176, further comprising utilizing a client.

20

288.   A method according to claim 287, wherein said client comprises at least one of the following:

end point software; end point hardware; tamper resistant software; tamper resistant hardware;

25   client side software; and client side hardware.

289.   A method according to claim 287, comprising utilizing said client for at least one of the following:

monitoring of client side storage; monitoring of client side access;

30   monitoring of client side usage;

monitoring of client side distribution;

monitoring of copying of information object excerpts;

monitoring of clipboard;

monitoring of at least one application;

monitoring of at least one interface;

control of at least one application;

control of at least one interface;

control of clipboard;

5   control of copying of information object excerpts;

control of client side storage;

control of client side access;

control of client side usage; and

control of client side distribution.

10

290.   A method according to claim 176, comprising utilizing comparing of at least two information objects to calculate pairwise similarity between objects.

15   291.   A method according to claim 290, comprising utilizing said pairwise similarity to map said information objects to a space.

292.   A method according to claim 291, wherein said space is an Euclidian space, and wherein the closeness between any two objects within said

20   Euclidian space is approximately proportional to said pairwise similarity between said information objects.

293.   A method according to claim 291, wherein said space is a weighted graph, and wherein the weight of an edge between any two objects

25   within said graph space is approximately proportional to said pairwise similarity between said information objects.

294.   A method according to claim 291, wherein said space is a graph, and wherein the existence of an edge between any two objects within said graph

30   space is dependent on said pairwise similarity between said information objects.

295.   A method according to claim 291, wherein said space is utilized to identify at least one similarity information class, wherein said information class consists of at least two information objects, wherein said information class

policy is a policy shared by the information class, and wherein said similarity information class is bounded within said space.

296.   A method according to claim 291, comprising utilizing said space to identify at last one information object substantially similar to an unidentified information object.

297.   A method according to claim 291, comprising using said space to identify at least one other information object substantially similar to an information object for which policy is not known, thereby to obtain a policy associated with said other information object to use as basis for a policy for said information object.

298.   A method according to claim 176, comprising storing information about said information object in a database.

299.   A method according to claim 176, further comprising extracting a descriptor of said information object, based on statistical analysis of said information object.

300.   A method according to claim 176, comprising storing the order of said elementary information units within said information object in a database.

301.   A method according to claim 300, comprising using said order for identification of said information object.

302.   A method according to claim 176, further comprising interfacing at least one of an information management system; and a document management system.

303.   A method according to claim 176, further comprising tracking at least one of the following:
usage patterns; storage patterns; and distribution patterns.

304.   A method according to claim 303, wherein said tracking is carried out to infer information about at least one of the following:

normal usage patterns; normal storage patterns; normal distribution patterns; irregular usage patterns; irregular storage patterns; and irregular distribution patterns.

305.   A method according to claim 304, wherein said inferred information is used to define at least part of a policy.

306.   A method according to claim 304, comprising using said inferred information for information object classification.

307.   A method according to claim 176, further comprising logging.

308.   A method according to claim 307, wherein said logging comprising logging of at least one of the following:

actions; events; and information objects identification.

309.   A method according to claim 307, wherein at least part of said logging is controlled by a policy.

310.   A method according to claim 307, wherein at least part of said logging is stored in a database.

311.   A method according to claim 307, comprising utilizing said logging to augment lifecycle information for said information object.

312.   A method according to claim 176, further comprising assessing the integrity of at least one information object, wherein said integrity assessment consists of comparing said information object with a version of said information object for which integrity is assured.

313.   A method according to claim 312, further comprising issuing a certificate of said integrity for at least one information object.

314.   A method according to claim 313, wherein said certificate is a

5   cryptographic certificate.

315.   A method according to claim 312, further comprising replacing said information object with said version of said information object for which said integrity is assured.

10

316.   A method according to claim 312, comprising identifying when said integrity of said information object is not satisfactory, and in such a case not allowing distribution of said information object.

15   317.   A method according to claim 312, comprising identifying when said integrity of said information object is not satisfactory, and in such a case not allowing storage of said information object.

318.   A method according to claim 312, comprising identifying when

20   said integrity of said information object is not satisfactory, and in such a case not allowing usage of said information object.

319.   A method according to claim 176, further comprising defining at least one constituent information object to be an ignored information object, and

25   wherein, whenever said to be ignored information object is an element of a compound information, ignoring said object in identification of said compound information object.

320.   A method according to claim 199, further comprising changing

30   access control information in accordance with said policy.

321.   A method according to claim 176, further comprising not allowing usage of respective ones of said information objects outside an organization.

322.   A method according to claim 176, further comprising not allowing storage of respective ones of said information object outside an organization.

323.   A method according to claim 176, further comprising not allowing distribution of respective ones of said information object outside an organization.

324.   A method according to claim 201, wherein said policy comprises at least one mandatory lifecycle.

325.   A method according to claim 324, wherein said action is dependent on the matching of said mandatory lifecycle with a lifecycle of a respective event.

326.   A method according to claim 324, wherein said mandatory lifecycle comprises at least one mandatory recipient of said information object; and an order of events concerning said information object.

327.   A method according to claim 205, wherein said inserting an additional part to said information object comprises inserting at least one of the following: a header; a footer; and a disclaimer.

328.   A method according to claim 199 or claim 213, further comprising defining areas and wherein said policy is dependent on whether an action is taken inside a given defined area.

329.   A method according to claim 199 or claim 213, further comprising defining areas and wherein said policy is dependent on whether an event occurs inside a given defined area.

330. A method according to claim 176 or claim 179, comprising using said deducing to locate at least one information object with similar content to a given information object.

5   331. A method according to claim 199, comprising attaching a respective policy to information objects according to their logical location within an information storage medium.

332. A method according to claim 331, further comprising utilizing a
10   crawler for automatic location of information objects within said information storage medium.

333. A method according to claim 332, wherein said information storage medium
15   is a file system.

334. A method according to claim 330, wherein said locating is done in an information storage medium.

20   335. A method according to claim 334, further comprising utilizing a crawler for automatic location of information objects within said information storage medium.

336. A method according to claim 334, wherein said information storage
25   medium comprises at least one file system.

337. A method for automated computerized exchange of information within an information object having overall coherency, the method comprising selecting amongst and carrying out at least one of the following:
30   deleting part of said information;
     replacing part of said information; and
     inserting an additional part to said information,
     wherein said carrying out additionally comprises preservation of the coherency of said information within said information object.

338.   A method according to claim 337, wherein said changing of said information is done in order to eliminate parts having policies that do not allow for at least one action to be executed while they are in the document.

5   339.   A method according to claim 337, wherein said changing of said information is carried out in order to personalize said information.

340.   A method according to claim 337, wherein said changing of said information is carried out in order to customize said information for a specific
10   use.

341.   A method according to claim 337, wherein said preserving said coherency comprises at least one of:

maintaining seamlessness; preserving the structure of said information;
15   preserving the linguistic coherency of said information; preserving the formatting style of said information; and preserve the pagination style of said information.

342.   A method according to claim 337, wherein said information
20   objects comprise compound information objects and wherein said changing of said information object is made to constituent parts of a compound information object.

343.   A method according to claim 337, carried out over a network
25   having users with different access rights to said information object, said selecting and carrying out being to adapt said information object to conform to access rights of a one of said users to whom said information object is released.

344.   Apparatus for automatic information identification to enforce an
30   information management policy on information objects, the apparatus comprising:

a scanning module for finding elementary information units within said information object; and

a deduction module for deducing information about the identity of said information object from identification of said elementary information units found within said information object, said deduced identity being usable to obtain a corresponding policy rule for applying to said information object.

345. Apparatus according to claim 344, wherein said information objects comprise at least one simple information object, said simple information object comprising one of the following:

an elementary information unit;

a set of elementary information units; and

an ordered set of elementary information units.

346. Apparatus according to claim 344, wherein said elementary information units comprise at least one of the following:

a sentence; a sequences of words; a word; a sequence of characters; a character; a sequence of numbers; a number; a sequence of digits; a digit; a vector; a curve; a pixel; a block of pixels; an audio frame; a musical note; a musical bar; a visual object; a sequence of video frames; a sequence of musical notes; a sequence of musical bars; and a video frame.

347. Apparatus according to claim 344, wherein said deduction module is further configured to assign elementary information unit identifiers to elementary information units after identification.

348. Apparatus according to claim 347, wherein said deduction module is further configured to utilize said elementary information unit identifiers in said deducing.

349. Apparatus according to claim 344, wherein said information object identification is carried out on an instance of said information object, said information object instance being said information object in a specific format.

350.   Apparatus according to claim 347, wherein said deduction module is configured to provide said elementary information unit identifiers in a manner determined at least partly by the content of said elementary information units which they are assigned to.

351.   Apparatus according to claim 350, wherein said elementary information unit identifiers are solely determined by said content.

352.   Apparatus according to claim 347, wherein said deduction module is configured to provide said elementary information units identifiers in a manner at least partly determined by locations within an information object of respective elementary information units to which they are assigned.

353.   Apparatus according to claim 344 further comprising a policy attachment unit associated with said deduction module, said policy attachment unit being configured to use said deducing to attach to said information object an information object policy, said policy comprising at least one of the following:

an allowed distribution of said information object;

a restriction on distribution of said information object;

an allowed storage of said information object;

a restriction on storage of said information object;

an action to be taken as a reaction to an event;

an allowed usage of said information object; and

a restriction on usage of said information object.

354.   Apparatus according to claim 344, wherein said deducing comprises utilizing conditional probabilities for at least one of the following:

identification of information objects;

classification of information objects; and

identification of a knowledge domain of information objects.

355.   Apparatus for automated computerized exchange of information within an information object having overall coherency, the apparatus comprising a selector for selecting amongst at least one of the following data modifications:

5   a deletion of part of said information;

a replacement of part of said information; and

an insertion of an additional part to said information,

the apparatus further comprising a data modification unit associated with said selector for carrying out said selected modification within said information

10   object, said data modification unit being associated with a coherency retention module for detecting coherency features of said information object and altering said modification in order to preserve said detected coherency features within said information object.

15   356.   Apparatus for automatic information identification of information objects, the apparatus comprising:

a scanning module for finding elementary information units within said information object; and

a deduction module for deducing information about the identity of said

20   information object from identification of said elementary information units found within said information object, said deduced identity being usable for controlling use of said information object.

357.   Apparatus according to claim 356, wherein said deduction

25   module is further configured to assign elementary information unit identifiers to elementary information units after identification.